PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/705, A61K 38/17, C07K 16/28, C12Q 1/68, G01N 33/68

A1

(11) International Publication Number: WO 98/49306

(43) International Publication Date: 5 November 1998 (05.11.98)

(21) International Application Number: PCT/US98/08791

(22) International Filing Date: 29 April 1998 (29.04.98)

(30) Priority Data:
08/846,523 29 April 1997 (29.04.97) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 08/846,523 (CIP)
Filed on 29 April 1997 (29.04.97)

(71) Applicant (*for all designated States except US*): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). SHAH, Purvi [IN/US]; 1608 Queen Charlotte Drive #5, Sunnyvale, CA 94087 (US).

(74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, GB, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: HUMAN C–TYPE LECTIN

(57) Abstract

The present invention provides a human C–type lectin (CTL–1) and polynucleotides which encode CTL–1. The invention also provides expression vectors, host cells, agonists, antisense molecules, antibodies or antagonists. The invention also provides methods for treating disorders associated with expression of CTL–1.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HUMAN C-TYPE LECTIN

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of a new human protein, CTL-1, and to the use of these sequences in the diagnosis, prevention, and treatment of disease.

BACKGROUND ART

Lectins are proteins which are defined by their ability to bind carbohydrates and agglutinate cells. Lectins have been shown to be involved in a wide variety of cellular functions including cell-cell and cell-matrix interactions. Lectins are widespread among plants, invertebrates and mammals.

Animal lectins have been grouped into four distinct families: 1) C-type lectins, which include selectins; 2) P-type lectins; 3) galectins (formerly termed S-type lectins or S-Lac lectins); and 4) pentraxins [Barondes SH *et al.* (1994) J. Biol. Chem. 269:20807-10]. The C-type lectins bind carbohydrate ligands in a Ca^{2+}-dependent manner and are structurally related to the asialoglycoprotein receptor. Selectins, a subcategory of the C-type lectins, are composite transmembrane molecules which are involved in cell-cell interactions. The selectins include lymphocyte homing receptors and platelet/endothelial cell surface receptors [Stoolman (1989) Cell 56:907-10].

C-type animal lectins contain Ca^{2+}-dependent carbohydrate-recognition domains (CRDs). The prototypical C-type animal lectins are integral membrane proteins (e.g., the asialoglycoprotein receptor); however, a number of soluble C-type animal lectins have been identified. The C-type lectins are placed in groups based upon their overall architecture, the position of the CRD relative to other domains, or upon the degree of homology between CRDs. Group V C-type lectins are type II transmembrane proteins (i.e., they have a cytoplasmic amino-terminus and an extracellular carboxy-terminus) which are found on the surface of natural killer (NK) lymphocytes as well as on T and B lymphocytes [Drickamer (1993) Curr. Opin. Struc. Biol. 3:393-400]. In mammals, NK and cytotoxic T cells eliminate tissues recognized as foreign, infected or compromised in some way [Gumperz and Parham (1995) Nature 378:245-48].

Members of the group V C-type lectin family include type II transmembrane proteins encoded by the NK gene complex which is located on chromosome 12 in humans and chromosome 6 in mice; the majority of these proteins exist as disulfide-linked homodimers on the surface of cells. Proteins encoded by the NK gene complex locus include NKR-P1, Ly49, NKG2, CD23, CD69 and CD94. The genes within the NK complex represent polymorphic multigene families. The NKR-P1 family of proteins trigger cellular activation, transduce stimulatory signals

and stimulate lysis of tumor cells when cross-linked. The Ly49 proteins function as class I MHC receptors and permit NK cells to recognize class I MHC molecules on target cells and to kill only cells lacking self MHC molecules. The CD69 protein (also called EA1, MLR3, Leu 23, AIM and BL-Ac/p26) functions as an activating molecule on lymphocytes (T, B and NK), platelets, neutrophils and eosinophils; data indicates that CD69 functions as a central molecule in cellular activation cascades [Hamann *et al.* (1993) J. Immunol. 150:4920-27]. A polymorphic locus encoding group V C-type lectins has been identified in the chicken [Bernot et al. (1994) Immunogenetics 39:221-29]. The chicken 17.5 gene is located within the Rfp-Y system which encodes class I and II MHC molecules. The 17.5 gene was found to be preferentially expressed in mature B cells. The homology between the chicken 17.5.3 protein and other NK receptors (NKR-P1, Ly49, NKG2) as well as the fact that these proteins are encoded by polymorphic multigene families suggests that the chicken 17.5 gene may encode an avian NK receptor.

Clearly the group V C-type lectins play important roles in cellular activation and proliferation of a variety of hematopoietic cells and in regulation of lysis of target cells, including tumor cells, by NK cells.

DISCLOSURE OF THE INVENTION

The invention features a new protein hereinafter designated human CTL-1 and characterized as having similarity to the human CD69 and chicken 17.5.3 proteins.

Accordingly, the invention features a substantially purified polypeptide having the amino acid sequence shown in SEQ ID NO:1 or fragments thereof. Preferred fragments of SEQ ID NO:1 are fragments of about 15 amino acids or greater in length which define fragments unique (i.e., having less than about 25% identity to fragments of another protein) to SEQ ID NO:1 or which retain biological activity or immunological activity (i.e., capable of eliciting anti-human CTL-1 antibodies). Fragments of SEQ ID NO:1 which are at least 25 amino acids, at least 50 amino acids, at least 100 amino acids, at least 125 amino acids, and at least 145 amino acids in length are also contemplated.

The invention further provides isolated and substantially purified polynucleotide sequences encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof. In a particular aspect, the polynucleotide is the nucleotide sequence of SEQ ID NO:2 or variants thereof. In another embodiment, the present invention provides polynucleotides comprising fragments of SEQ ID NO:2 having a length greater than 5 nucleotides. The invention further contemplates fragments of this polynucleotide sequence (*i.e.*, SEQ ID NO:2) that are at least 6 nucleotides, at least 50 nucleotides, at least 100 nucleotides, at

least 250 nucleotides, at least 500 nucleotides and at least 650 nucleotides in length.

In addition, the invention provides polynucleotide sequences which hybridize under stringent conditions to the polynucleotide sequence of SEQ ID NO:2. In another embodiment, the present invention provides a composition comprising an isolated and purified polynucleotide sequence encoding human CTL-1.

The invention provides polynucleotide sequences comprising the complement of SEQ ID NO:2 or variants thereof; these complementary nucleic acid sequences may comprise the complement of the entire nucleic acid sequence of SEQ ID NO:2 or fragments thereof. In another embodiment, the present invention provides a composition comprising an isolated and purified polynucleotide sequence comprising the complement of SEQ ID NO:2 or variants thereof.

The invention additionally features nucleic acid sequences encoding fragments, portions or the complement of the polynucleotides encoding CTL-1.

The invention also provides an isolated polynucleotide comprising at least a portion of the nucleic acid sequence of SEQ ID NO:2 or variants thereof contained on a recombinant expression vector. In yet another aspect, the expression vector containing the polynucleotide sequence is contained within a host cell. The invention is not limited by the nature of the host cell employed. For example, the host cell may be an E. coli cell, a yeast cell, an insect cell, a mammalian cell, etc.

The present invention also provides a method for producing a polypeptide comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof, the method comprising the steps of: a) culturing the host cell containing an expression vector containing an isolated polynucleotide encoding at least a fragment of the human CTL-1 polypeptide under conditions suitable for the expression of the polypeptide; and b) recovering the polypeptide from the host cell culture.

In another embodiment, the invention provides a pharmaceutical composition comprising a substantially purified human CTL-1 protein having the amino acid sequence of SEQ ID NO:1 in conjunction with a suitable pharmaceutical carrier.

The invention also provides a purified antibody which binds specifically to a polypeptide comprising at least a portion of the amino acid sequence of SEQ ID NO:1.

The invention provides a purified agonist which specifically binds to and modulates the activity of a polypeptide comprising at least a portion of the amino acid sequence of SEQ ID NO:1. The present invention further provides a pharmaceutical composition comprising a purified agonist which specifically binds to and modulates the activity of a polypeptide comprising at least a portion of the amino acid sequence of SEQ ID NO:1.

The invention also provides a purified antagonist which specifically binds to and modulates the activity of a polypeptide comprising at least a portion of the amino acid sequence of SEQ ID NO:1. Still further, the invention further provides a pharmaceutical composition comprising a purified antagonist which specifically binds to and modulates the activity of a polypeptide comprising at least a portion of the amino acid sequence of SEQ ID NO:1.

The invention also provides a method for treating cancer comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a purified antagonist which modulates the activity of a polypeptide comprising at least a portion of the amino acid sequence of SEQ ID NO:1. The treatment of a variety of cancers (i.e., tumors), including but not limited to ovary, paraganglionic, penis, brain, thyroid and heart tumors, using agonists as well as antagonists of human CTL-1 is also contemplated by the present invention.

The invention also provides a method for the detection of polynucleotides encoding human CTL-1 in a biological sample comprising the steps of: a) hybridizing a polynucleotide sequence encoding human CTL-1 (SEQ ID NO:1) to nucleic acid material of a biological sample, thereby forming a hybridization complex; and b) detecting the hybridization complex, wherein the presence of the complex correlates with the presence of a polynucleotide encoding human CTL-1 in the biological sample. In a preferred embodiment, prior to hybridization, the nucleic acid material of the biological sample is amplified by the polymerase chain reaction. In another preferred embodiment, the nucleic acid material comprises metaphase chromosomes prepared from human cells (e.g., from a biopsy or blood sample) and detection of the hybridization complex indicates the chromosomal location (i.e., normal or rearranged) of the CTL-1 gene in the human cells.

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A and 1B show the amino acid sequence (SEQ ID NO:1) and nucleic acid sequence (SEQ ID NO:2) of human CTL-1. The alignment was produced using MacDNASIS PRO™ software (Hitachi Software Engineering Co., Ltd., San Bruno, CA).

Figures 2A and 2B show the amino acid sequence alignments between the human CTL-1 (SEQ ID NO:1), human CD69 (GI 291898; SEQ ID NO:3) and chicken 17.5.3 (GI 505325; SEQ ID NO:4) proteins. The alignment was produced using the multisequence alignment program of DNASTAR™ software (DNASTAR Inc, Madison WI).

Figure 3 shows the hydrophobicity plot (MacDNASIS PRO software) for human CTL-1 (SEQ ID NO:1); the positive X axis reflects amino acid position, and the negative Y axis,

hydrophobicity.

Figures 4A and B show the northern analysis for SEQ ID NO:2. The northern analysis was produced electronically using LIFESEQ™ database (Incyte Pharmaceuticals, Inc., Palo Alto, CA).

MODES FOR CARRYING OUT THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors, and reagents described as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are reported in the publications which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. Similarly, "amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules.

A "composition comprising a given polynucleotide sequence" as used herein refers broadly to any composition containing the given polynucleotide sequence. The composition may

comprise an aqueous solution. Compositions comprising polynucleotide sequences encoding human CTL-1 (SEQ ID NO:1) or fragments thereof (e.g., SEQ ID NO:2 and fragments thereof) may be employed as hybridization probes. In this case, the human CTL-1-encoding polynucleotide sequences are typically employed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., SDS) and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms, such as "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

"Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al. (1993) Anticancer Drug Des. 8:53-63).

CTL-1, as used herein, refers to the amino acid sequences of substantially purified CTL-1 obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic, or recombinant.

"Consensus", as used herein, refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, or which has been extended using XL-PCR™ (Perkin Elmer, Norwalk, CT) in the 5' and/or the 3' direction and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte clone using the GELVIEW™ Fragment Assembly system (GCG, Madison, WI), or which has been both extended and assembled.

A "variant" of human CTL-1, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNASTAR software. Furthermore, as described herein, certain amino acid residues which are highly conserved among group V C-type

lectins are located within the CRD of these C-type lectins. It is preferred that these conserved residues not be substituted (as will be understood by those in the art, conservative substitutions may be tolerated at some of the conserved residues), inserted or deleted when producing variants of human CTL-1.

A "deletion", as used herein, refers to a change in either amino acid or nucleotide sequence in which one or more amino acid or nucleotide residues, respectively, are absent.

An "insertion" or "addition", as used herein, refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid or nucleotide residues, respectively, as compared to the naturally occurring molecule.

A "substitution", as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

The term "biologically active", as used herein, refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic human CTL-1, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "agonist", as used herein, refers to a molecule which, when bound to human CTL-1, causes a change in human CTL-1 which modulates the activity of human CTL-1. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to human CTL-1.

The terms "antagonist" or "inhibitor", as used herein, refer to a molecule which, when bound to human CTL-1, blocks or modulates the biological or immunological activity of human CTL-1. Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to human CTL-1.

The term "modulate", as used herein, refers to a change or an alteration in the biological activity of human CTL-1. Modulation may be an increase or a decrease in protein activity, a change in binding characteristics, or any other change in the biological, functional, or immunological properties of human CTL-1.

The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of human CTL-1 or portions thereof and, as such, is able to effect some or all of the actions of human CTL-1-like molecules.

The term "derivative", as used herein, refers to the chemical modification of a nucleic acid encoding human CTL-1 or the encoded human CTL-1. Illustrative of such modifications

would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of the natural molecule.

The term "substantially purified", as used herein, refers to nucleic or amino acid sequences that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

"Amplification" as used herein refers to the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies well known in the art (Dieffenbach, C.W. and G.S. Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY).

The term "hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex", as used herein, refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen binds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., membranes, filters, chips, pins or glass slides to which cells have been fixed for in situ hybridization).

The terms "complementary" or "complementarity", as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, for the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be "partial", in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands.

The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic

acid; it is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

As known in the art, numerous equivalent conditions may be employed to comprise either low or high stringency conditions. Factors such as the length and nature (DNA, RNA, base composition) of the sequence, nature of the target (DNA, RNA, base composition, presence in solution or immobilization, etc.), and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency different from, but equivalent to, the above listed conditions.

The term "stringent conditions", as used herein, is the "stringency" which occurs within a range from about Tm-5°C (5°C below the melting temperature (Tm) of the probe) to about 20°C to 25°C below Tm. As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect identical or related polynucleotide sequences. Under "stringent conditions" SEQ ID NO:2 or fragments thereof will hybridize to its exact complement and closely related sequences. The stringent conditions are chosen such that SEQ ID NO:2 or fragments thereof will hybridize to sequences encoding human CTL-1 but not to sequences encoding human CD69 (GI 291897; SEQ ID NO:5 or its RNA equivalents) or the chicken 17.5.3 protein (GI 505324; SEQ ID NO:6 or its RNA equivalents). When fragments of SEQ ID NO:2 are employed in hybridization reactions, the stringent conditions include the choice of fragments of SEQ ID NO:2 to be used. Fragments of SEQ ID NO:2 which contain unique sequences (*i.e.*, regions which are either non-homologous to or which contain less than about 50% homology or complementarity with SEQ ID NOS:5 or 6) are preferentially employed.

The term "antisense", as used herein, refers to nucleotide sequences which are

complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in reference to the sense strand.

The term "portion", as used herein, with regard to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein "comprising at least a portion of the amino acid sequence of SEQ ID NO:1" encompasses the full-length human CTL-1 and fragments thereof.

"Transformation", as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being transformed and may include, but is not limited to, viral infection, electroporation, lipofection, and particle bombardment. Such "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells which transiently express the inserted DNA or RNA for limited periods of time.

The term "antigenic determinant", as used herein, refers to that portion of a molecule that makes contact with a particular antibody (i.e., an epitope). When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to a given region or three-dimensional structure on the protein; these regions or structures are referred to as antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The terms "specific binding" or "specifically binding", as used herein, in reference to the interaction of an antibody and a protein or peptide, mean that the interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) on the protein; in

other words, the antibody is recognizing and binding to a specific protein structure rather than to proteins in general. For example, if an antibody is specific for epitope "A", the presence of a protein containing epitope A (or free, unlabeled A) in a reaction containing labeled "A" and the antibody will reduce the amount of labeled A bound to the antibody.

The term "sample", as used herein, is used in its broadest sense. A biological sample suspected of containing nucleic acid encoding human CTL-1 or fragments thereof may comprise a cell, chromosomes isolated from a cell (e.g., a spread of metaphase chromosomes), genomic DNA (in solution or bound to a solid support such as for Southern analysis), RNA (in solution or bound to a solid support such as for northern analysis), cDNA (in solution or bound to a solid support), an extract from cells or a tissue, and the like.

The term "correlates with expression of a polynucleotide", as used herein, indicates that the detection of the presence of ribonucleic acid that is similar to SEQ ID NO:2 by northern analysis is indicative of the presence of mRNA encoding human CTL-1 in a sample and thereby correlates with expression of the transcript from the polynucleotide encoding the protein.

"Alterations" in the polynucleotide of SEQ ID NO:2, as used herein, comprise any alteration in the sequence of polynucleotides encoding human CTL-1 including deletions, insertions, and point mutations that may be detected using hybridization assays. Included within this definition is the detection of alterations to the genomic DNA sequence which encodes human CTL-1 (e.g., by alterations in the pattern of restriction fragment length polymorphisms capable of hybridizing to SEQ ID NO:2), the inability of a selected fragment of SEQ ID NO: 2 to hybridize to a sample of genomic DNA (e.g., using allele-specific oligonucleotide probes), and improper or unexpected hybridization, such as hybridization to a locus other than the normal chromosomal locus for the polynucleotide sequence encoding human CTL-1 (e.g., using fluorescent in situ hybridization [FISH] to metaphase chromosomes spreads).

As used herein, the term "antibody" refers to intact molecules as well as fragments thereof, such as Fa, $F(ab')_2$, and Fv, which are capable of binding the epitopic determinant. Antibodies that bind human CTL-1 polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest as the immunizing antigen. The polypeptide or peptide used to immunize an animal can be derived from the transition of RNA or synthesized chemically, and can be conjugated to a carrier protein, if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin and thyroglobulin. The coupled peptide is then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

The term "humanized antibody", as used herein, refers to antibody molecules in which

amino acids have been replaced in the non-antigen binding regions in order to more closely resemble a human antibody, while still retaining the original binding ability.

THE INVENTION

The invention is based on the discovery of a new human protein (CTL-1), the polynucleotides encoding human CTL-1, and the use of these compositions for the diagnosis, prevention, or treatment of diseases associated with altered or abnormal human CTL-1 expression. As mRNA encoding human CTL-1 is found in a number of tumors, human CTL-1 serves as a marker for cancerous cells, particularly ovary, paraganglionic, penis, brain, thyroid and heart tumor cells.

Nucleic acids encoding the human CTL-1 of the present invention were first identified in Incyte Clone 1756224 from the PITUNOT03 cDNA library through a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 1756224/ PITUNOT03 and 2599243/ UTRSNOT010.

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1, as shown in Figures 1A and 1B. Human CTL-1 contains 149 amino acids, a number of which are residues shown to be conserved among the CRD of C-type animal lectins; residues 35-149 represent the CRD of human CTL-1. A conserved sequence motif found in C-type CRDs is described by Drickamer (supra) and a version of this motif is found in the PROSITE database as the C-type lectin domain signature (CTL). Sequences corresponding to the CTL within the human CTL-1 protein of the present invention include, but are not limited to, G_{33}, C_{63}, E_{78}, W_{92}, I_{93}, G_{94}, L_{95}, G_{108}, C_{123}, G_{130}, C_{136} and C_{144}. Human CTL-1 contains one major hydrophobic segment or transmembrane domain which is located at residues 8-25 of SEQ ID NO:1. CTL-1 lacks a typical secretory signal which suggests that the amino-terminus of CTL-1, like that of other group V C-type lectins, is intracytoplasmic.

CTL-1 contains eight cysteine residues (i.e., C_8, C_{35}, C_{46}, C_{63}, C_{87}, C_{123} C_{136} and C_{144}). In addition to providing sites for disulfide bond formation, the cysteine residues provide potential sites for palmitoylation. Seven of the eight cysteine residues found in human CTL-1 are conserved in location with cysteine residues found in the human CD69 and the chicken 17.5.3 proteins (i.e., C_8, C_{35}, C_{46}, C_{63}, C_{123} C_{136} and C_{144} of CTL-1) (see Fig. 2). The human CTL-1 of the present invention contains numerous potential O-linked glycosylation sites (i.e., serine and threonine residues). CTL-1 has three potential N-linked glycoslyation sites (i.e., Asn-Xaa-

Ser/Thr) (*i.e.*, N_{57}, N_{62} and N_{100}). In addition, CTL-1 contains numerous potential phosphorylation sites (i.e., typically the hydroxyl groups of serine, threonine and tyrosine residues although asparagine, histidine and lysine residues may also be phosphorylated). In particular, CTL-1 has three potential protein kinase C phosphorylation sites (Ser/Thr-Xaa-Arg/Lys) located at residues 58-60, 133-135 and 138-140, two potential casein kinase II phosphorylation sites (i.e., residues 50-53 and 71-74) and one potential tyrosine kinase phosphorylation site (i.e., residues 118-125).

CTL-1 has chemical and structural homology with the human CD69 protein (GI 291898; SEQ ID NO:3) [Hamann et al., supra] and the chicken 17.5.3 protein (GI 505325; SEQ ID NO:4) [Bernot et al., supra]. In particular, over their respective CRDs (residues 35-149 of CTL-1; residues 85-199 of CD69 and residues 129-257 of the chicken 17.5.3 protein), human CTL-1 and human CD69 share 35.6% identity and 55% similarity and human CTL-1 and the chicken 17.5.3 protein share 45% identity. A pair of residues are said to be similar if they represent conservative substitutions. Figure 2 provides an alignment between the amino acid sequences of SEQ ID NOS:1, 3 and 5.

CTL-1 shares chemical and structural homology with other group V C-type lectins including, but not limited to, mouse CD69. As discussed above, group V C-type lectins are associated with cellular activation and proliferation and the regulation of the lysis of target cells, including tumor cells.

The hydrophobicity plot of the human CTL-1 protein is provided in Figure 3. The isoelectric point of human CTL-1 is 8.82.

Northern analysis (Fig. 4) shows the expression of CTL-1-encoding sequences in various libraries, at least 29% of which are cancerous and at least 18% of which are involved with the hematopoietic system and/or immune response, including inflammatory and/or autoimmune disease (e.g., rheumatoid synovium). Of particular note is the expression of CTL-1 mRNA in ovarian tumor, paraganglionic tumor, penis tumor, brain tumor, adenomatous goiter of the thyroid and heart tumor (myoma) libraries. This pattern of expression demonstrates that CTL-1 serves as a marker for cancerous cells, particularly ovarian tumor cells. In addition, Fig. 4 demonstrates that CTL-1 mRNA is expressed in activated hematopoietic cells (e.g., adherent and non-adherent cells from mixed lymphocyte reactions and IL-5 treated mononuclear cells). This pattern of expression suggests that human CTL-1 expression is upregulated during cellular activation as are other members of the group V C-type lectin family such as the early activation antigen CD69 (Hamann et al., supra). Fig. 4 further demonstrates that at least 25% of the libraries containing

human CTL-1 sequences are derived from brain and/or neural tissue (normal as well as tumorous) and thus CTL-1 serves as a marker of brain and neural tissues.

The invention also encompasses CTL-1 variants. A preferred CTL-1 variant is one having at least 80%, and more preferably 90%, amino acid sequence identity to the human CTL-1 amino acid sequence (SEQ ID NO:1). A most preferred CTL-1 variant is one having at least 95% amino acid sequence identity to SEQ ID NO:1.

The invention also encompasses polynucleotides which encode human CTL-1. Accordingly, any nucleic acid sequence which encodes the amino acid sequence of human CTL-1 can be used to generate recombinant molecules which express CTL-1. In a particular embodiment, the invention encompasses the polynucleotide comprising the nucleic acid sequence of SEQ ID NO:2 as shown in Figures 1A and 1B.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding human CTL-1, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring CTL-1, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode human CTL-1 and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring human CTL-1 under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding human CTL-1 or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding human CTL-1 and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences, or portions thereof, which encode human CTL-1 and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression

vectors and cell systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding human CTL-1 or any portion thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those shown in SEQ ID NO:2, under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency.

Altered nucleic acid sequences encoding human CTL-1 which are encompassed by the invention include deletions, insertions, or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent human CTL-1. The encoded protein may also contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent CTL-1. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of human CTL-1 is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; phenylalanine and tyrosine.

Also included within the scope of the present invention are alleles of the genes encoding CTL-1. As used herein, an "allele" or "allelic sequence" is an alternative form of the gene which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

Methods for DNA sequencing which are well known and generally available in the art may be used to practice any embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham,

Chicago, IL), or combinations of recombinant polymerases and proofreading exonucleases such as the ELONGASE Amplification System marketed by Gibco/BRL (Gaithersburg, MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) and the ABI 377 DNA sequencers (Perkin Elmer).

The nucleic acid sequences encoding CTL-1 may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). In particular, genomic DNA is first amplified in the presence of primer to linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). The primers may be designed using OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may also be used to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before performing PCR.

Another method which may be used to retrieve unknown sequences is that of Parker, J.D. et al. (1991; Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk in genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been

size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into the 5' and 3' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled devise camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g. Genotyper™ and Sequence Navigator™, Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode human CTL-1, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of CTL-1 in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express human CTL-1.

As will be understood by those of skill in the art, it may be advantageous to produce human CTL-1-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter human CTL-1 encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences.

For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding human CTL-1 may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of human CTL-1 activity, it may be useful to encode a chimeric human CTL-1 protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the human CTL-1 encoding sequence and the heterologous protein sequence, so that human CTL-1 may be cleaved and purified away from the heterologous moiety.

In another embodiment, sequences encoding human CTL-1 may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of human CTL-1, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J.Y. et al. (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer).

The newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, supra). Additionally, the amino acid sequence of CTL-1, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a biologically active human CTL-1, the nucleotide sequences encoding human CTL-1 or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence.

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding human CTL-1 and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in

Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, NY, and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding CTL-1. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript® phagemid (Stratagene, LaJolla, CA) or pSport1™ plasmid (Gibco BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO; and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding human CTL-1, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for CTL-1. For example, when large quantities of CTL-1 are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional E. coli cloning and expression vectors such as Bluescript® (Stratagene), in which the sequence encoding CTL-1 may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of ß-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors (Promega, Madison, WI) may also be used to express foreign polypeptides as

fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) Methods Enzymol. 153:516-544. In cases where plant expression vectors are used, the expression of sequences encoding CTL-1 may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L.E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, NY; pp. 191-196.

An insect system may also be used to express human CTL-1. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding human CTL-1 may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of human CTL-1 will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which human CTL-1 may be expressed (Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. 91:3224-3227).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding human CTL-1 may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing human CTL-1 in

infected host cells (Logan, J. and Shenk, T. (1984) Proc. Natl. Acad. Sci. 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding human CTL-1. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding human CTL-1, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of he inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells available from American Type Culture Collection (ATCC) Rockville, MD such as CHO (ATCC CCL 61 and CRL 9618), HeLa (ATCC CCL 2), MDCK (ATCC CCL 34 and CRL 6253), HEK 293 (ATCC CRL 1573), WI-38 (ATCC CCL 75), which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express CTL-1 may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using

tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1980) Cell 22:817-23) genes which can be employed in tk^- or $aprt^-$ cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. et al (1981) J. Mol. Biol. 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, ß glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. et al. (1995) Methods Mol. Biol. 55:121-131). .

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding human CTL-1 is inserted within a marker gene sequence, recombinant cells containing sequences encoding human CTL-1 can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding human CTL-1 under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain the nucleic acid sequence encoding human CTL-1 and express human CTL-1 may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

The presence of polynucleotide sequences encoding human CTL-1 can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or portions or fragments of polynucleotides encoding human CTL-1. Nucleic acid amplification based assays involve the use

of oligonucleotides or oligomers based on the sequences encoding human CTL-1 to detect transformants containing DNA or RNA encoding human CTL-1. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

A variety of protocols for detecting and measuring the expression of human CTL-1, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbant assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on human CTL-1 is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. et al. (1983; J. Exp. Med. 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding human CTL-1 include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding human CTL-1, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); and U.S. Biochemical Corp., Cleveland, OH). Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding CTL-1 may be cultured under conditions suitable for the expression and recovery of the protein from cell culture, including from cell membrane preparations. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode human CTL-1 may be designed to contain signal sequences which direct secretion of

portions of human CTL-1 through a prokaryotic or eukaryotic cell membrane. When it is desired to express a secreted form of CTL-1, a polynucleotide sequence encoding the extracellular domain(s) (i.e., residues 26-149 of SEQ ID NO:1) is preferentially employed. Alternatively, CTL-1 may be expressed as a membrane-bound protein in a host cell and the recombinant CTL-1 recovered from the membrane of the host cell using techniques well known to the art .

Other recombinant constructions may be used to join sequences encoding human CTL-1 to nucleotide sequences encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and human CTL-1 may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing human CTL-1 and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography as described in Porath, J. et al. (1992, Prot. Exp. Purif. 3: 263-281) while the enterokinase cleavage site provides a means for purifying human CTL-1 from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. et al. (1993; DNA Cell Biol. 12:441-453).

In addition to recombinant production, fragments of human CTL-1 may be produced by direct peptide synthesis using solid-phase techniques Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of human CTL-1 may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

THERAPEUTICS

Based on the chemical and structural homology among CTL-1 (SEQ ID NO:1) and the human CD69 and chicken 17.5.3 proteins (SEQ ID NOS:3 and 4) and other proteins in the group V C-type lectin family, CTL-1 appears to be a new group V C-type lectin. Human CD69 has been shown to be rapidly expressed following activation of lymphocytes. Following expression of CD69 on the cell surface, binding of anti-CD69 antibodies induces the proliferation of T and B cells. Stimulation of CD69 also induces platelet aggregation and degranulation of neutrophils.

Stimulation of CD69 is associated with an influx of Ca^{2+} and the activation of protein kinase C, events which are associated with cellular proliferation in a wide variety of cell types.

Based on the homology between CTL-1 and human CD69 protein, as well as the homology between CTL-1 and other members of the group V C-type lectin family, it is believed that CTL-1 plays a role in the regulation of lysis of target cells, including tumor cells, by NK cells and in the regulation of cell growth. The expression pattern of CTL-1 (Fig. 4) suggests that CTL-1 expression is associated with cellular activation. Improper regulation of cell growth is observed in tumors and thus, the amino acid and nucleic acid sequences of CTL-1 provided herein provide a means of producing therapeutic compounds for the treatment of disease states associated with altered CTL-1 expression.

Increased expression of CTL-1 may lead to uncontrolled cell growth and cancer. Therefore, in one embodiment, antagonists of CTL-1, including anti-CTL-1 antibodies, may be administered to a subject to treat and/or prevent cancer, including adenocarcinoma, sarcoma, melanoma, lymphoma, leukemia, ganglioneuroma, and myeloma. In particular, types of cancer may include, but are not limited to, cancer of the ovaries, paraganglion, penis, brain, lung thyroid, heart, uterus, blood, small intestine, and spleen.

In another embodiment, antagonists of CTL-1 may be administered in combination with other conventional chemotherapeutic agents. The combination of therapeutic agents having different mechanisms of action will have synergistic effects allowing for the use of lower effective doses of each agent and lessening side effects.

In another embodiment, a vector expressing the complement of the polynucleotide encoding CTL-1 or an antisense molecule may be administered to a subject to treat or prevent any of the types of cancer listed above.

As several group V C-type lectins are associated with NK-directed lysis of tumor cells, in certain cases increasing the activity of human CTL-1 in a patient's NK cells, by increasing the density of CTL-1 on the NK cell surface and/or by the addition of agonists of CTL-1, may be useful for increasing the elimination of cancer cells in a patient. Therefore, in another embodiment, agonists of CTL-1 may be administered in combination with other conventional chemotherapeutic agents to treat cancer. The combination of therapeutic agents having different mechanisms of action will have synergistic effects allowing for the use of lower effective doses of each agent and lessening side effects.

Antibodies which are specific for human CTL-1 may be used directly as an antagonist, or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or

tissue which express CTL-1. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies, (i.e., those which reduce or abolish CTL-1 activity) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with CTL-1 or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the peptides, fragments, or oligopeptides used to induce antibodies to human CTL-1 have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids. It is also preferable that they are identical to a portion of the amino acid sequence of the natural protein, and they may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of CTL-1 amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule.

Monoclonal antibodies to human CTL-1 may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce CTL-1-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be

generated by chain shuffling from random combinatorial immunoglobin libraries (Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:11120-3).

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for human CTL-1 may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between human CTL-1 and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering CTL-1 epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides encoding human CTL-1, or any fragment thereof, or antisense molecules, may be used for therapeutic purposes. In one aspect, antisense to the polynucleotide encoding human CTL-1 may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding human CTL-1. Thus, antisense molecules may be used to modulate CTL-1 activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding human CTL-1.

Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express antisense molecules complementary to the polynucleotides of the gene encoding CTL-1. These techniques are described both in

Sambrook et al. (supra) and in Ausubel et al. (supra).

Genes encoding human CTL-1 can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide or fragment thereof which encodes CTL-1. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA, or PNA, to the control regions of the gene encoding CTL-1, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples which may be used include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding CTL-1.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for

chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding human CTL-1. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods which are well known in the art.

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of human CTL-1, antibodies to human CTL-1, mimetics, agonists, antagonists, or inhibitors of human CTL-1. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial,

intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of CTL-1, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example CTL-1 or fragments thereof, antibodies of CTL-1, agonists, antagonists or inhibitors of CTL-1, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind CTL-1 may be used for the diagnosis of conditions or diseases characterized by expression of CTL-1, or in assays to monitor patients being treated with CTL-1, agonists, antagonists or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for CTL-1 include methods which utilize the antibody and a label

to detect CTL-1 in human cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring CTL-1 are known in the art and provide a basis for diagnosing altered or abnormal levels of CTL-1 expression. Normal or standard values for CTL-1 expression are established by combining cell extracts (e.g., extracts comprising membrane proteins) or tissue sections taken from normal mammalian subjects, preferably human, with antibody to CTL-1 under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric, means. Quantities of CTL-1 expressed in subject, control and disease, samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding CTL-1 are used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of CTL-1 may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of CTL-1, and to monitor regulation of CTL-1 levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding CTL-1 or closely related molecules, may be used to identify nucleic acid sequences which encode CTL-1. The specificity of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding CTL-1, alleles, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably contain at least 50% of the nucleotides from any of the CTL-1 encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of SEQ ID NO:2 or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring CTL-1.

Means for producing specific hybridization probes for DNAs encoding CTL-1 include the cloning of nucleic acid sequences encoding CTL-1 or CTL-1 derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S, or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding CTL-1 may be used for the diagnosis of conditions or diseases which are associated with expression of CTL-1. Examples of such conditions or diseases include ovaries, paraganglion, penis, brain, lung thyroid, heart, uterus, blood, small intestine, and spleen. The polynucleotide sequences encoding CTL-1 may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered CTL-1 expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding CTL-1 provide the basis for assays that detect activation or induction of various cancers, particularly those mentioned above; in addition the lack of expression of CTL-1 may be detected using the CTL-1-encoding nucleotide sequences disclosed herein. The nucleotide sequences encoding CTL-1 may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequences have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding CTL-1 in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of disease associated with expression of CTL-1, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, which encodes CTL-1, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a

substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease.

Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or metastasis, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding CTL-1 may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5'->3') and another with antisense (3'<-5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of CTL-1 include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (Melby, P.C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In another embodiment of the invention, the nucleic acid sequences which encode CTL-1 may also be used to generate hybridization probes which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH,

FACS, or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C.M. (1993) Blood Rev. 7:127-134, and Trask, B.J. (1991) Trends Genet. 7:149-154.

FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, NY) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding CTL-1 on a physical chromosomal map and a specific disease, or predisposition to a specific disease, may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier, or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R.A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals. As a number of genes encoding group V C-type lectins localize to the NK gene complex on human chromosome 12 (e.g., NKG2, NKR-P1A and CD94), sequences encoding CTL-1 are presumed to map to human chromosome 12. Thus, CTL-1 sequences serve as a marker for human chromosome 12.

In another embodiment of the invention, CTL-1, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between CTL-1 and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to CTL-1 large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with CTL-1, or fragments thereof, and washed. Bound CTL-1 is then detected by methods well known in the art. Purified CTL-1 can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding CTL-1 specifically compete with a test compound for binding CTL-1. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with CTL-1.

In additional embodiments, the nucleotide sequences which encode CTL-1 may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

INDUSTRIAL APPLICABILITY

I PITUNOT03 cDNA Library Construction

The PITUNOT03 cDNA library was constructed from microscopically normal tissue obtained from a 46-year-old Caucasian male (specimen #RA95-05-0337; International Institute for the Advancement of Medicine, Exton, PA) who had died from colon cancer.

The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury, NJ) in guanidinium isothiocyanate solution. The lysates were extracted once with acid phenol at pH 4.0 per Stratagene's RNA isolation protocol (Stratagene, Inc., San Diego, CA). The RNA was extracted twice with an equal volume of acid phenol, reprecipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in DEPC-treated water and DNase treated for 25 min at 37°C. mRNAs were isolated with the Qiagen Oligotex kit (QIAGEN, Inc.; Chatworth, CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (Cat. #18248-013; Gibco/BRL,

Gaithersburg, MD). cDNAs were fractionated on a Sepharose CL4B column (Cat. #275105-01; Pharmacia), and those cDNAs exceeding 400 bp were ligated into pSport I (Gibco/BRL). The plasmid pSport I was subsequently transformed into DH5α™ competent cells (Cat. #18258-012; Gibco/BRL).

II Isolation, and Sequencing of cDNA clones

Plasmid DNA was released from the cells and purified using the REAL Prep 96 Plasmid Kit (Catalog #26173; QIAGEN, Inc.). This kit enabled the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Catalog #22711,Gibco/BRL) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the cultures were incubated for 19 hours and at the end of incubation, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4° C.

The cDNAs were sequenced by the method of Sanger et al. (1975, J. Mol. Biol. 94:441f), using a Hamilton Micro Lab 2200 (Hamilton, Reno, NV) in combination with Peltier Thermal Cyclers (PTC200 from MJ Research, Watertown, MA) and Applied Biosystems 377 DNA Sequencing Systems.

III Homology Searching of cDNA Clones and Their Deduced Proteins

The nucleotide sequences of the Sequence Listing or amino acid sequences deduced from them were used as query sequences against databases such as GenBank, SwissProt, BLOCKS, and Pima II. These databases which contain previously identified and annotated sequences were searched for regions of homology (similarity) using BLAST, which stands for Basic Local Alignment Search Tool (Altschul, S.F. (1993) J. Mol. Evol. 36:290-300; Altschul et al. (1990) J. Mol. Biol. 215:403-410).

BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal or plant) origin. Other algorithms such as the one described in Smith RF and TF Smith (1992; Protein Engineering 5:35-51), incorporated herein by reference, can be used when dealing with primary sequence patterns and secondary structure gap penalties. As disclosed in this application, the sequences have lengths of at least 49 nucleotides, and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach, as detailed in Karlin, S. and S.F. Atschul (1993; Proc. Nat. Acad. Sci. 90:5873-7) and incorporated herein by reference, searches for matches between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. In this application, threshold was set at 10^{-25} for nucleotides and 10^{-14} for peptides.

Incyte nucleotide sequences were searched against the GenBank databases for primate (pri), rodent (rod), and mammalian sequences (mam), and deduced amino acid sequences from the same clones are searched against GenBank functional protein databases, mammalian (mamp), vertebrate (vrtp) and eukaryote (eukp), for homology. The relevant database for a particular match were reported as a GIxxx±p (where xxx is pri, rod, etc and if present, p = peptide).

A comparison of the full-length and partial cDNA sequences and the deduced amino acid sequences corresponding to the human CTL-1 gene and CTL-1 protein with known nucleotide and protein sequences in GenBank revealed that the human CTL-1 protein and nucleotide sequences shared some homology with the human CD69 protein and cDNA sequences (SEQ ID NOS:3 and 5, respectively) and the chicken 17.5.3 protein and cDNA sequences (SEQ ID NOS:4 and 6, respectively) as well as the mouse CD69 protein (GI 584907: SEQ ID NO:7).

IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook et al., supra).

Analogous computer techniques using BLAST (Altschul, S.F. 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database (Incyte Pharmaceuticals). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analysis are reported as a list of libraries in which the transcript encoding CTL-1 occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

Electronic northern analysis (Fig. 4) revealed that mRNA encoding human CTL-1 (SEQ ID NO:1) was present in libraries generated from a variety of adult and fetal tissues. In addition to expression in apparently normal human tissues, CTL-1 was expressed in a variety of tumors, including ovary, paraganglionic, penis, brain, thyroid and heart tumors. CTL-1 mRNA was expressed in activated hematopoietic cells (e.g., adherent and non-adherent cells from mixed lymphocyte reactions and IL-5 treated mononuclear cells). This pattern of expression suggests that human CTL-1 expression is upregulated during cellular activation.

V Extension of CTL-1-Encoding Polynucleotides

Incyte Clone 1756224 or SEQ ID NO:2 is used to design oligonucleotide primers for extending a partial nucleotide sequence to full length or for obtaining 5' or 3', intron or other control sequences from genomic libraries. One primer is synthesized to initiate extension in the antisense direction (XLR) and the other is synthesized to extend sequence in the sense direction (XLF). Primers are used to facilitate the extension of the known sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest. The initial primers are designed from the cDNA using OLIGO 4.06 (National Biosciences), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations is avoided.

The original, selected cDNA libraries, or a human genomic library are used to extend the sequence; the latter is most useful to obtain 5' upstream regions. If more extension is necessary or desired, additional sets of primers are designed to further extend the known region.

By following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 40 pmol of

each primer and the recommended concentrations of all other components of the kit, PCR is performed using the Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA) and the following parameters:

Step 1 94° C for 1 min (initial denaturation)
Step 2 65° C for 1 min
Step 3 68° C for 6 min
Step 4 94° C for 15 sec
Step 5 65° C for 1 min
Step 6 68° C for 7 min
Step 7 Repeat step 4-6 for 15 additional cycles
Step 8 94° C for 15 sec
Step 9 65° C for 1 min
Step 10 68° C for 7:15 min
Step 11 Repeat step 8-10 for 12 cycles
Step 12 72° C for 8 min
Step 13 4° C (and holding)

A 5-10 μl aliquot of the reaction mixture is analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products are selected and removed from the gel. Further purification involves using a commercial gel extraction method such as QIAQuick™ (QIAGEN Inc., Chatsworth, CA). After recovery of the DNA, Klenow enzyme is used to trim single-stranded, nucleotide overhangs creating blunt ends which facilitate religation and cloning.

After ethanol precipitation, the products are redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase are added, and the mixture is incubated at room temperature for 2-3 hours or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) are transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium (Sambrook et al., supra). After incubation for one hour at 37° C, the whole transformation mixture is plated on Luria Bertani (LB)-agar (Sambrook et al., supra) containing 2x Carb. The following day, several colonies are randomly picked from each plate and cultured in 150 μl of liquid LB/2x Carb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture is transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 μl of each sample is transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for

the extension reaction are added to each well. Amplification is performed using the following conditions:

Step 1 94° C for 60 sec
Step 2 94° C for 20 sec
Step 3 55° C for 30 sec
Step 4 72° C for 90 sec
Step 5 Repeat steps 2-4 for an additional 29 cycles
Step 6 72° C for 180 sec
Step 7 4° C (and holding)

Aliquots of the PCR reactions are run on agarose gels together with molecular weight markers. The sizes of the PCR products are compared to the original partial cDNAs, and appropriate clones are selected, ligated into plasmid, and sequenced.

VI Labeling and Use of Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham) and T4 polynucleotide kinase (DuPont NEN®, Boston, MA). The labeled oligonucleotides are substantially purified with Sephadex G-25 superfine resin column (Pharmacia & Upjohn). A portion containing 10^7 counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II; DuPont NEN®).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham, NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester, NY) is exposed to the blots in a Phosphoimager cassette (Molecular Dynamics, Sunnyvale, CA) for several hours, hybridization patterns are compared visually.

VII Complementary Sequences

Sequences complementary to the CTL-1-encoding sequence, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring CTL-1. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same

procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequences of CTL-1, as shown in Figures 1A and 1B, is used to inhibit expression of naturally occurring CTL-1. The complementary oligonucleotide is designed from the most unique 5' sequence as shown in Figures 1A and 1B and used either to inhibit transcription by preventing promoter binding to the upstream nontranslated sequence or translation of an CTL-1-encoding transcript by preventing the ribosome from binding. Using an appropriate portion of the signal and 5' sequence of SEQ ID NO:2, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or 5' coding sequence of the polypeptide as shown in Figures 1A and 1B.

VIII Expression of CTL-1

Expression of CTL-1 is accomplished by subcloning the cDNAs into appropriate vectors and transforming the vectors into host cells. In this case, the cloning vector, pSport (Gibco/BRL) is used to express CTL-1 or fragments thereof in E. coli. Upstream of the cloning site, this vector contains a promoter for ß-galactosidase, followed by sequence containing the amino-terminal Met, and the subsequent seven residues of ß-galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

Induction of an isolated, transformed bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first eight residues of ß-galactosidase, about 5 to 15 residues of linker, and the full length protein or fragments thereof. Sequences encoding the extracellular domain of CTL-1 located at residues 26-149 of SEQ ID NO:1 are preferentially employed for the production of soluble forms of recombinant CTL-1, including CTL-1 fusion proteins. The signal residues present on the pSport vector direct the secretion of CTL-1 polypeptides into the bacterial growth media.

Alternatively, CTL-1 may be expressed as a membrane-bound protein in a host cell and the recombinant CTL-1 recovered from the membrane of the host cell using techniques well known to the art.

IX Demonstration of CTL-1 Activity

Given the chemical and structural similarity between the human CTL-1 and other group V C-type lectins, human CTL-1 is presumed to be a C-type lectin and therefore capable of binding to carbohydrates. The ability of recombinant human CTL-1 to bind carbohydrates may be demonstrated by examining the ability of human CTL-1 to bind to affinity columns comprising carbohydrates (e.g., lactose, maltose, D-mannose, D-galactose, etc. which are available from

Sigma Chemical Corp., St. Louis, MO) or by using the assay described by Christa *et al.* (1994) FEBS Lett. 337:114]. Preferably a soluble form of CTL-1 (i.e., lacking the transmembrane domain) is expressed and applied to the affinity column.

Some C-type lectins are known to agglutinate bacteria. The ability of human CTL-1 to agglutinate bacteria is examined using the assay described by Iovanna *et al.* [(1991) J. Biol. Chem. 266:24664]. Briefly, bacteria (e.g., E. coli strains KH802 or JM101) are grown at 37°C to stationary phase in L-broth. The bacteria are then collected by centrifugation and washed in PBS. The washed bacteria are resuspended in PBS containing 0.5 mM $CaCl_2$ (PBS/$CaCl_2$) and are placed in the wells of microtiter plates at a concentration of approximately 5 x 10^7 bacteria/200 μl PBS/$CaCl_2$. Human CTL-1(preferably a soluble form of CTL-1) is then added at a variety of concentrations (e.g., 1 to 50 μg/ml) and the presence of macroscopic aggregation is monitored following a 3 hour incubation at 25°C. Concanavalin A and albumin at 50 μg/ml may be employed as positive and negative controls, respectively.

X Production of CTL-1 Specific Antibodies

Human CTL-1 that is substantially purified using PAGE electrophoresis (Sambrook, supra), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence deduced from SEQ ID NO:2 is analyzed using DNASTAR software (DNASTAR Inc) to determine regions of high immunogenicity and a corresponding oligopolypeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions, is described by Ausubel et al. (supra), and others.

Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma, St. Louis, MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel et al., supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

XI Purification of Naturally Occurring CTL-1 Using Specific Antibodies

Naturally occurring or recombinant human CTL-1 is substantially purified by immunoaffinity chromatography using antibodies specific for CTL-1. An immunoaffinity column is constructed by covalently coupling CTL-1 antibody to an activated chromatographic resin, such as CnBr-activated Sepharose (Pharmacia & Upjohn). After the coupling, the resin is

blocked and washed according to the manufacturer's instructions.

Media containing human CTL-1 is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of human CTL-1 (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/CTL-1 binding (eg, a buffer of pH 2-3 or a high concentration of a chaotrope, such as urea or thiocyanate ion), and CTL-1 is collected.

XII Identification of Molecules Which Interact with CTL-1

Human CTL-1 or biologically active fragments thereof are labeled with ^{125}I Bolton-Hunter reagent (Bolton et al. (1973) Biochem. J. 133: 529). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled CTL-1, washed and any wells with labeled CTL-1 complex are assayed. Data obtained using different concentrations of CTL-1 are used to calculate values for the number, affinity, and association of CTL-1 with the candidate molecules.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: INCYTE PHARMACEUTICALS, INC.

(ii) TITLE OF THE INVENTION: HUMAN C-TYPE LECTIN

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) PCT APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/846,523
(B) FILING DATE: 29-APR-197

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0281 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 650-855-0555
(B) TELEFAX: 650-845-4166
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 149 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: PITUNOT03
(B) CLONE: 1756224

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Met Thr Lys His Lys Lys Cys Phe Ile Ile Val Gly Val Leu Ile
 1               5                  10                  15
Thr Thr Asn Ile Ile Thr Leu Ile Val Lys Leu Thr Arg Asp Ser Gln
            20                  25                  30
Ser Leu Cys Pro Tyr Asp Trp Ile Gly Phe Gln Asn Lys Cys Tyr Tyr
        35                  40                  45
Phe Ser Lys Glu Glu Gly Asp Trp Asn Ser Ser Lys Tyr Asn Cys Ser
    50                  55                  60
Thr Gln His Ala Asp Leu Thr Ile Ile Asp Asn Ile Glu Glu Met Asn
65                  70                  75                  80
Phe Leu Arg Arg Tyr Lys Cys Ser Ser Asp His Trp Ile Gly Leu Lys
                85                  90                  95
Met Ala Lys Asn Arg Thr Gly Gln Trp Val Asp Gly Ala Thr Phe Thr
            100                 105                 110
Lys Ser Phe Gly Met Arg Gly Ser Glu Gly Cys Ala Tyr Leu Ser Asp
        115                 120                 125
Asp Gly Ala Ala Thr Ala Arg Cys Tyr Thr Glu Arg Lys Trp Ile Cys
    130                 135                 140
Arg Lys Arg Ile His
145
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 655 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
- (A) LIBRARY: PITUNOT03
- (B) CLONE: 1756224

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
CTTCAAAAGT GGTATACCTC TAGTTTGGAG CTGTGCTGTA AAAACAAGAG TAACATTTTT     60
ATATTAAAGT TAAATAAAGT TACAACTTTG AAGAGAGTTT CTGCAAGACA TGACACAAAG    120
CTGCTAGCAG AAAATCAAAA CGCTGATTAA AAGAAGCACG GTATGATGAC CAAACATAAA    180
AAGTGTTTTA TAATTGTTGG TGTTTTAATA ACAACTAATA TTATTACTCT GATAGTTAAA    240
CTAACTCGAG ATTCTCAGAG TTTATGCCCC TATGATTGGA TTGGTTTCCA AAACAAATGC    300
TATTATTTCT CTAAAGAAGA AGGAGATTGG AATTCAAGTA AATACAACTG TTCCACTCAA    360
CATGCCGACC TAACTATAAT TGACAACATA GAAGAAATGA ATTTTCTTAG GCGGTATAAA    420
TGCAGTTCTG ATCACTGGAT TGGACTGAAG ATGGCAAAAA ATCGAACAGG ACAATGGGTA    480
GATGGAGCTA CATTTACCAA ATCGTTTGGC ATGAGAGGGA GTGAAGGATG TGCCTACCTC    540
AGCGATGATG GTGCAGCAAC AGCTAGATGT TACACCGAAA GAAAATGGAT TTGCAGGAAA    600
AGAATACACT AAGTTAATGT CTAAGATAAT GGGGAAAATA GAAAATAACA TTATT         655
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 199 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
- (A) LIBRARY: GenBank
- (B) CLONE: 291898

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Ser Ser Glu Asn Cys Phe Val Ala Glu Asn Ser Ser Leu His Pro
 1               5                  10                  15
Glu Ser Gly Gln Glu Asn Asp Ala Thr Ser Pro His Phe Ser Thr Arg
            20                  25                  30
His Glu Gly Ser Phe Gln Val Pro Val Leu Cys Ala Val Met Asn Val
        35                  40                  45
Val Phe Ile Thr Ile Leu Ile Ile Ala Leu Ile Ala Leu Ser Val Gly
    50                  55                  60
Gln Tyr Asn Cys Pro Gly Gln Tyr Thr Phe Ser Met Pro Ser Asp Ser
65                  70                  75                  80
His Val Ser Ser Cys Ser Glu Asp Trp Val Gly Tyr Gln Arg Lys Cys
                85                  90                  95
Tyr Phe Ile Ser Thr Val Lys Arg Ser Trp Thr Ser Ala Gln Asn Ala
            100                 105                 110
Cys Ser Glu His Gly Ala Thr Leu Ala Val Ile Asp Ser Glu Lys Asp
        115                 120                 125
Met Asn Phe Leu Lys Arg Tyr Ala Gly Arg Glu Glu His Trp Val Gly
    130                 135                 140
Leu Lys Lys Glu Pro Gly His Pro Trp Lys Trp Ser Asn Gly Lys Glu
145                 150                 155                 160
Phe Asn Asn Trp Phe Asn Val Thr Gly Ser Asp Lys Cys Val Phe Leu
                165                 170                 175
Lys Asn Thr Glu Val Ser Ser Met Glu Cys Glu Lys Asn Leu Tyr Trp
            180                 185                 190
Ile Cys Asn Lys Pro Tyr Lys
        195
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 257 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 505325

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Pro Leu Phe Leu Ser Phe Ser Pro Arg Ser Leu Arg Glu Val Leu
 1               5                  10                  15
Ala Lys Lys Ser Ala Pro Pro Ala Pro Leu Cys Pro Gln Pro Asp Pro
            20                  25                  30
Ser Leu Leu Leu Ser Leu His Ala Ala Gly Ala Val Pro His Leu Tyr
        35                  40                  45
Asp Ala Thr Glu Glu Lys Glu Arg Leu Ser Pro Ser Pro Pro Arg Glu
    50                  55                  60
Ala Thr Thr Arg Glu Gly Asp Glu Glu Arg Gln Ser Gln Arg Gly Ser
65                  70                  75                  80
Gly Cys Ser Glu Leu Arg Gln Asn Arg Arg Arg Val Leu Cys Val Ala
                85                  90                  95
Leu Ser Ala Val Pro Cys Met Leu Val Leu Ala Leu Val Ala Val Ile
            100                 105                 110
Val Leu Gln Arg Pro Ser Cys Ser Pro Arg Pro Pro Phe Ser His Val
        115                 120                 125
Cys Pro Asn Ala Trp Val Gly Phe Gln Gly Lys Cys Tyr Tyr Phe Ser
    130                 135                 140
Asp Thr Glu Ser Asp Trp Asn Ser Ser Arg Glu His Cys His Arg Leu
145                 150                 155                 160
```

```
Gly Ala Ser Leu Ala Thr Leu Asp Thr Lys Glu Glu Met Glu Phe Met
                165                 170                 175
Leu Gln Tyr Gln Arg Pro Ala Asp Arg Trp Ile Gly Leu His Arg Ala
            180                 185                 190
Glu Gly Asp Glu His Trp Thr Trp Ala Asp Gly Ser Ala Phe Thr Asn
        195                 200                 205
Arg Pro Val Phe Glu Leu Arg Gly Gly Gly Arg Cys Ala Tyr Leu Asn
    210                 215                 220
Gly Asp Gly Ile Ser Ser Ala Leu Cys His Ser Glu Lys Phe Trp Val
225                 230                 235                 240
Cys Ser Arg Ala Asp Ser Tyr Val Arg Trp Arg Lys Gly Thr Asn Pro
                245                 250                 255
Gln
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 1662 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
  (A) LIBRARY: GenBank
  (B) CLONE: 291897

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CAAGAGCTCC AGCAAAGACT TTCACTGTAG CTTGACTTGA CCTGAGATTA ACTAGGGAAT     60
CTTGAGAATA AAGATGAGCT CTGAAAATTG TTTCGTAGCA GAGAACAGCT CTTTGCATCC    120
GGAGAGTGGA CAAGAAAATG ATGCCACCAG TCCCCATTTC TCAACACGTC ATGAAGGGTC    180
CTTCCAAGTT CCTGTCCTGT GTGCTGTAAT GAATGTGGTC TTCATCACCA TTTTAATCAT    240
AGCTCTCATT GCCTTATCAG TGGGCCAATA CAATTGTCCA GGCCAATACA CATTCTCAAT    300
GCCATCAGAC AGCCATGTTT CTTCATGCTC TGAGGACTGG GTTGGCTACC AGAGGAAATG    360
CTACTTTATT TCTACTGTGA AGAGGAGCTG GACTTCAGCC CAAAATGCTT GTTCTGAACA    420
TGGTGCTACT CTTGCTGTCA TTGATTCTGA AAAGGACATG AACTTTCTAA AACGATACGC    480
AGGTAGAGAG GAACACTGGG TTGGACTGAA AAAGGAACCT GGTCACCCAT GGAAGTGGTC    540
AAATGGCAAA GAATTTAACA ACTGGTTCAA CGTTACAGGG TCTGACAAGT GTGTTTTTCT    600
GAAAAACACA GAGGTCAGCA GCATGGAATG TGAGAAGAAT TTATACTGGA TATGTAACAA    660
ACCTTACAAA TAATAAGGAA ACGTGTTCAC TTATTGACTA TTATAGAATG GAACTCAAGG    720
AAATCTGTGT CAGTGGATGC TGCTCTGTGG TCCGAAGTCT TCCATAGAGA CTTTGTGAAA    780
AAAAATTTTA TAGTGTCTTG GGAATTTTCT TCCAAACAGA ACTATGGAAA AAAAGGAAGA    840
AATTCCAGGA AAATCTGCAC TTGTGGCTTT TATTGCCATG AGCTAGAAGC ATCACAGGTT    900
GACCAATAAC CATGCCCAAG AATGAGAAGA ATGACTATGC AACCTTTGGA TGCACTTTAT    960
ATTATTTTGA ATCCAGAAAT AATGAAATAA CTAGGCGTGG ACTTACTATT AATTGCTGAA   1020
TGACTACCAA CAGTGAGAGC CCTTCATGCA TTTGCACTAT TGGAAGGAGT TAGATGTTGG   1080
TACTAGATAC TGAATGTAAA CAAAGGAATT ATGGCTGGTA ACATAGTTTT TAGTCTAATT   1140
GAATCCCTTA AACTCAGGGA GCATTTATAA ATGGCAAATG CTTATGAAAC TAAGATTTGT   1200
AATATTTCTC TCTTTTTAGA GAAATTTGCC AATTTACTTT GTTATTTTTC CCCAAAAAGA   1260
ATGGGATGAT CGTGTATTTA TTTTTTTACT TCCTCAGCTG TAGACAGGTC CTTTTCGATG   1320
GTACATATTT CTTTGCCTTT ATAATCTTTT ATACAGTGTC TTACAGAGAA AAGACATAAG   1380
CAAAGACTAT GAGGAATATT TGCAAGACAT AGAATAGTGT TGGAAAATGT GCAATATGTG   1440
ATGTGGCAAA TCTCTATTAG GAAATATTCT GTAATCTTCA GACCTAGAAT AATACTAGTC   1500
TTATAATAGG TTTGTGACTT TCCTAAATCA ATTCTATTAC GTGCAATACT TCAATACTTC   1560
ATTTAAAATA TTTTTATGTG CAATAAAATG TATTTGTTTG TATTTTGTGT TCAGTACAAT   1620
TATAAGCTGT TTTTATATAT GTGAAATAAA AGTAGAATAA AC                       1662
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

Met	Asp	Ser	Glu	Asn	Cys	Ser	Ile	Thr	Glu	Asn	Ser	Ser	Ser	His	Leu
1				5					10					15	
Glu	Arg	Gly	Gln	Lys	Asp	His	Gly	Thr	Ser	Ile	His	Phe	Glu	Lys	His
			20					25					30		
His	Glu	Gly	Ser	Ile	Gln	Val	Ser	Ile	Pro	Trp	Ala	Val	Leu	Ile	Val
		35					40					45			
Val	Leu	Ile	Thr	Ser	Leu	Ile	Ile	Ala	Leu	Ile	Ala	Leu	Asn	Val	Gly
	50					55					60				
Lys	Tyr	Asn	Cys	Pro	Gly	Leu	Tyr	Glu	Lys	Leu	Glu	Ser	Ser	Asp	His
65					70					75					80
His	Val	Ala	Thr	Cys	Lys	Asn	Glu	Trp	Ile	Ser	Tyr	Lys	Arg	Thr	Cys
				85					90					95	
Tyr	Phe	Phe	Ser	Thr	Thr	Thr	Lys	Ser	Trp	Ala	Leu	Ala	Gln	Arg	Ser
			100					105					110		
Cys	Ser	Glu	Asp	Ala	Ala	Thr	Leu	Ala	Val	Ile	Asp	Ser	Glu	Lys	Asp
		115					120					125			
Met	Thr	Phe	Leu	Lys	Arg	Tyr	Ser	Gly	Glu	Leu	Glu	His	Trp	Ile	Gly
	130					135					140				
Leu	Lys	Asn	Glu	Ala	Asn	Gln	Thr	Trp	Lys	Trp	Ala	Asn	Gly	Lys	Glu
145					150					155					160
Phe	Asn	Ser	Trp	Phe	Asn	Leu	Thr	Gly	Ser	Gly	Arg	Cys	Val	Ser	Val
				165					170					175	
Asn	His	Lys	Asn	Val	Thr	Ala	Val	Asp	Cys	Glu	Ala	Asn	Phe	His	Trp
			180					185					190		
Val	Cys	Ser	Lys	Pro	Ser	Arg									
		195													

CLAIMS

What is claimed is:

1. A substantially purified human CTL-1 protein comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof.

2. An isolated and purified polynucleotide sequence encoding the human CTL-1 protein of claim 1.

3. A polynucleotide sequence which hybridizes under stringent conditions to the polynucleotide sequence of claim 2.

4. A composition comprising the polynucleotide sequence of claim 2.

5. An isolated and purified polynucleotide sequence comprising SEQ ID NO:2 or variants thereof.

6. A polynucleotide sequence which is complementary to SEQ ID NO:2 or variants thereof.

7. A composition comprising the polynucleotide sequence of claim 6.

8. An expression vector containing the polynucleotide sequence of claim 2.

9. A host cell containing the vector of claim 8.

10. A method for producing a polypeptide comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof, the method comprising the steps of:

a) culturing the host cell of claim 9 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

11. A pharmaceutical composition comprising a substantially purified human CTL-1 protein having an amino acid sequence of SEQ ID NO:1 in conjunction with a suitable pharmaceutical carrier.

12. A purified antibody which binds specifically to the polypeptide of claim 1.

13. A purified agonist which modulates the activity of the polypeptide of claim 1.

14. A purified antagonist which decreases the activity of the polypeptide of claim 1.

15. A method for treating cancer comprising administering to a subject in need of such treatment an effective amount of the purified antagonist of claim 14.

16. A method for detection of polynucleotides encoding human CTL-1 protein in a biological sample comprising the steps of:

a) hybridizing the polynucleotide of claim 2 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b) detecting said hybridization complex, wherein the presence of said complex correlates with the presence of a polynucleotide encoding human CTL-1 protein in said biological sample.

17. The method of claim 16, wherein before hybridization, the nucleic acid material of the biological sample is amplified by the polymerase chain reaction.

```
                9                    18                  27                   36                   45                   54
CTT CAA AAG TGG TAT ACC TCT AGT TTG GAG CTG TGC TGT AAA AAC AAG AGT AAC

               63                    72                  81                   90                   99                  108
ATT TTT ATA TTA AAG TTA AAT AAA GTT ACA ACT TTG AAG AGA GTT TCT GCA AGA

              117                   126                 135                  144                  153                  162
CAT GAC ACA AAG CTG CTA GCA GAA AAT CAA AAC GCT GAT TAA AAG AAG CAC GGT

              171                   180                 189                  198                  207                  216
ATG ATG ACC AAA CAT AAA AAG TGT TTT ATA ATT GTT GGT GTT TTA ATA ACA ACT
M   M   T   K   H   K   K   C   F   I   I   V   G   V   L   I   T   T

              225                   234                 243                  252                  261                  270
AAT ATT ATT ACT CTG ATA GTT AAA CTA ACT CGA GAT TCT CAG AGT TTA TGC CCC
N   I   I   T   L   I   V   K   L   T   R   D   S   Q   S   L   C   P

              279                   288                 297                  306                  315                  324
TAT GAT TGG ATT GGT TTC CAA AAC AAA TGC TAT TAT TTC TCT AAA GAA GAA GGA
Y   D   W   I   G   F   Q   N   K   C   Y   Y   F   S   K   E   E   G

              333                   342                 351                  360                  369                  378
GAT TGG AAT TCA AGT AAA TAC AAC TGT TCC ACT CAA CAT GCC GAC CTA ACT ATA
D   W   N   S   S   K   Y   N   C   S   T   Q   H   A   D   L   T   I
```

FIGURE 1A

```
        387         396         405         414         423         432
ATT GAC AAC ATA GAA GAA ATG AAT TTT CTT AGG CGG TAT AAA TGC AGT TCT GAT
I   D   N   I   E   E   M   N   F   L   R   R   Y   K   C   S   S   D

        441         450         459         468         477         486
CAC TGG ATT GGA CTG AAG ATG GCA AAA AAT CGA ACA GGA CAA TGG GTA GAT GGA
H   W   I   G   L   K   M   A   K   N   R   T   G   Q   W   V   D   G

        495         504         513         522         531         540
GCT ACA TTT ACC AAA TCG TTT GGC ATG AGA GGG AGT GAA GGA TGT GCC TAC CTC
A   T   F   T   K   S   F   G   M   R   G   S   E   G   C   A   Y   L

        549         558         567         576         585         594
AGC GAT GAT GGT GCA GCA ACA GCT AGA TGT TAC ACC GAA AGA AAA TGG ATT TGC
S   D   D   G   A   A   T   A   R   C   Y   T   E   R   K   W   I   C

        603         612         621         630         639         648
AGG AAA AGA ATA CAC TAA GTT AAT GTC TAA GAT AAT GGG GAA AAT AGA AAA TAA
R   K   R   I   H

CAT TAT T
```

FIGURE 1B

```
1    M - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   CTL-1
1    M S - - - - - - - - - - - - - - - - - - - - - S E N C F V A   g291898
1    M P L F L S F S P R S L R E V L A K K S A P P A P L C P Q P   g505325

2    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   CTL-1
10   E N S - - - S L H P - - - - - - - - E S G Q E N D - - - A T   g291898
31   D P S L L L S L H A A G A V P H L Y D A T E E K E R L S P S   g505325

2    - - - - - - - - - - - - - - - - - - - - - - - - M T K H K K   CTL-1
26   S P H F S T R H E G S - - - - - - - - - - - - - - - - F Q V   g291898
61   P P R E A T T R E G D E E R Q S Q R G S G C S E L R Q N R R   g505325

8    - - - C F I I V G V - - L I T T N I I T L I - V K L T R D S   CTL-1
40   P V L C A V M N V V - - F I T I L I I A L I A L S V G Q Y N   g291898
91   R V L C V A L S A V P C M L V L A L V A V I - V - L Q R P S   g505325

32   - - - - - - - Q - - - - - - - S L C P Y D W I G F Q N K C Y   CTL-1
68   C P G Q Y T F S M P S D S H V S S C S E D W V G Y Q R K C Y   g291898
119  C S P R P P F S - - - - - - - H V C P N A W V G F Q G K C Y   g505325
```

FIGURE 2A

```
48   YFSKEEGDWNSSKYNCSTQHADLTIIDNIE  CTL-1
98   FISTVKRSWTSAQNACSEHGATLAVIDSEK  g291898
142  YFSDTESDWNSSREHCHRLGASLATLDTKE  g505325

78   EMNFLRRYKCSSDHWIGLKMAKNRTG-QWV  CTL-1
128  DMNFLKRYAGREEHWVGLKKEPGHP-WKWS  g291898
172  EMEFMLQYQRPADRWIGLHRAEGDEHWTWA  g505325

107  DGATFTKS--FGMRGSEGCAYLSDDGAATA  CTL-1
157  NGKEFNNW--FNVTGSDKCVFLKNTEVSSM  g291898
202  DGSAFTNRPVFELRGGGRCAYLNGDGISSA  g505325

135  RCYTERKWICRK------------RIH     CTL-1
185  ECEKNLYWICNK--PY----------K     g291898
232  LCHSEKFWVCSRADSYVRWRKGTNPQ      g505325
```

FIGURE 2B

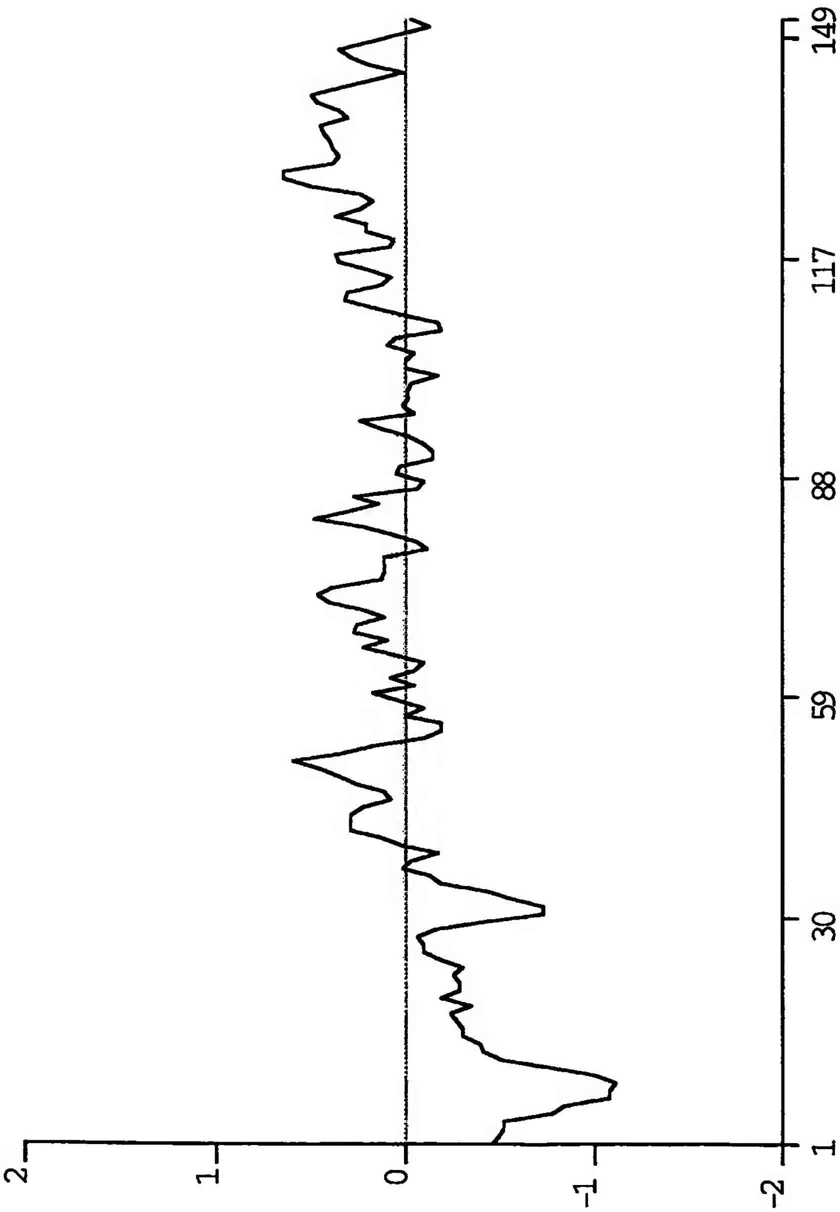

FIGURE 3

Library	Lib Description	Abun	Pct Abun
OVARTUM02	ovarian tumor, 64 F, WM	1	0.1002
PGANNON02	paraganglionic tumor, benign paraganglioma, 46 M, NORM	1	0.0780
PLACNOM01	placenta, fetal M, WM	1	0.0580
UTRSNOT12	uterus, myometrium, 41 F, match to UTRSTUT05	2	0.0568
PENITUT01	penis tumor, carcinoma, 64 M	2	0.0533
PROSNOT02	prostate, 50 M, match to PROSTUT01	1	0.0435
COLNNOT09	colon, 60 M, match to COLNTUT16	1	0.0390
PITUNOT03	pituitary, 46 M	1	0.0348
LUNGNOT18	lung, 66 F	1	0.0298
UTRSNOT10	uterus, endometrium, 50 F	1	0.0292
DRGCNOT01	ganglion, dorsal root, cervical, 32 M	1	0.0278
BRAITUT22	brain tumor, frontal parietal, 76 F	1	0.0271
LUNGNOT10	lung, fetal M	1	0.0261
THYRNOT09	thyroid, adenomatous goiter, 18 F	1	0.0254
TMLR3DT02	lymphocytes (non-adher PBMNC), M/F, 72-hr MLR	1	0.0246
HIPOAZT01	brain, hippocampus, Alzheimer's, 74 M	1	0.0245
SINJNOT02	small intestine, jejunum, 8 F	1	0.0241
MMLR1DT01	macrophages (adher PBMNC), M/F, 24-hr MLR	1	0.0236
OVARTUT03	ovarian tumor, carcinoma, 52 F	1	0.0235
OVARNOT02	ovary, 59 F	2	0.0225

FIGURE4A

Library	Lib Description	Abun	Pct Abun
SPLNNOT02	spleen, 29 M	1	0.0220
SYNORAB01	synovium, hip, rheumatoid, 68 F	1	0.0195
CERVNOT01	cervix, 35 F	1	0.0194
LATRTUT02	heart tumor, myoma, 43 M	1	0.0137
DRGLNOT01	ganglion, dorsal root, thoracic/lumbar, 32 M	1	0.0123
UCMCL5T01	mononuclear cells, treated IL-5	1	0.0084
BRAINOM01	brain, infant F, NORM, WM	1	0.0045

Electronic Northern Results returned a total of 27 row(s).

FIGURE4B

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/08791

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 A61K38/17 C07K16/28 C12Q1/68
G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HAMANN J. ET AL.: "AICL: a new activation-induced antigen encoded by the human NK gene complex." IMMUNOGENETICS, vol. 45, no. 5, March 1997, pages 295-300, XP002073213 see the whole document ---	1-17
X	EMBL database entry HS808161; Accession number H11808; 3. July 1995; Hillier et al.: 'The WashU-Merck EST project.' XP002073415 see abstract ---	1-11,16, 17

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 August 1998	13/08/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Mandl, B

1

Form PCT/ISA/210 (second sheet) (July 1992)

International application No
PCT/US 98/08791

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 98 21328 A (KATO SEISHI ;PROTEGENE INC (JP); SEKINE SHINGO (JP); SAGAMI CHEM R) 22 May 1998 see SEQ.IDs. 11, 36 and 61 ---	1-11,16, 17
P,X	WO 97 40069 A (GENETICS INST) 30 October 1997 see page 31, line 29 - page 32, line 9 -----	1-11,16, 17

1

INTERNATIONAL SEARCH REPORT

.ternational application No.
PCT/US 98/08791

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim(s) 15
is(are) directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 98/08791

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9821328 A	22-05-1998	NONE	
WO 9740069 A	30-10-1997	AU 2459397 A	12-11-1997
		AU 2728697 A	12-11-1997
		WO 9740151 A	30-10-1997